



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/734,866	12/12/2003	William V. Da Palma	BOC9-2003-0096 (1082-7U)	1522
46322	7590	02/05/2010	EXAMINER	
CAREY, RODRIGUEZ, GREENBERG & PAUL, LLP			COLUCCI, MICHAEL C	
STEVEN M. GREENBERG			ART UNIT	PAPER NUMBER
950 PENINSULA CORPORATE CIRCLE			2626	
SUITE 3020				
BOCA RATON, FL 33487				
		MAIL DATE	DELIVERY MODE	
		02/05/2010	PAPER	

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents  
United States Patent and Trademark Office  
P.O. Box 1450  
Alexandria, VA 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

**BEFORE THE BOARD OF PATENT APPEALS  
AND INTERFERENCES**

Application Number: 10/734,866

Filing Date: December 12, 2003

Appellant(s): DA PALMA ET AL.

01/26/2010

William DA PALMA et al.

For Appellant

**EXAMINER'S ANSWER**

This is in response to the appeal brief filed 12/14/2009 appealing from the Office action mailed 07/15/2009.

**(1) Real Party in Interest**

A statement identifying by name the real party in interest is contained in the brief.

**(2) Related Appeals and Interferences**

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

**(3) Status of Claims**

The statement of the status of claims contained in the brief is correct.

**(4) Status of Amendments After Final**

No amendment after final has been filed.

**(5) Summary of Claimed Subject Matter**

The summary of claimed subject matter contained in the brief is correct.

**(6) Grounds of Rejection to be Reviewed on Appeal**

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

**(7) Claims Appendix**

The copy of the appealed claims contained in the Appendix to the brief is correct.

**(8) Evidence Relied Upon**

US 20030156706	William et al.	11-2003
US 20030212561	Koehler et al.	08-2003

**(9) Grounds of Rejection**

The following ground(s) of rejection are applicable to the appealed claims:

***Claim Rejections - 35 USC § 103***

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. Claims 1, 8, 15, and 21-37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Williams et al. US 20030212561 A1 (hereinafter Williams) in view of

Koehler et al. US 20030156706 A1 (hereinafter Koehler).

Re claims 1, 8, and 15, Williams teaches a method for simulating ([0034]) a run-time user interaction with a voice application ([0047] & Fig. 6), said method comprising the steps of:

loading a user simulation script programmed to specify simulated voice interactions with the voice application ([0047] & Fig. 6);  
processing the user simulation script ([0034]) to generate both a simulated output for the voice application corresponding to the nominal output and a simulated input for the voice application corresponding to a pre-determined user input to the voice application (an expected voice input and a voice-recognized text input associated with a branch of the voice capable markup language application, and analyzing the expected voice input and the voice-recognized text input to determine an input to the tag, Williams [0070]).

However, Williams fails to teach deriving from the voice application a nominal output of the voice application.

Koehler teaches a user application, for example a simulator module 230 causes options to be provided to the training terminal 130 at step s312. The options may be provided textually or audibly (Koehler [0052]). Further, in Fig. 4, Koehler teaches pre-determined responses in a voice application such, wherein voice files may be referenced, prerecorded and stored or produced on-line by the developer, for each of the desired customer lines in a dialog segment. The data for the new scenario is then

stored in the database 212 as text. However, in an embodiment of the invention, the customer lines are converted to speech (unless they are already recorded speech) and also stored as audio files in .WAV or MP3 format, for example (Koehler [0087] & Fig. 4 element 414 for example).

Koehler teaches actual and simulated data outputs, wherein Koehler teaches that simulator module 230 creates a realistic training environment in two respects. First, through the business rules application 232, the simulator module 230 initiates a real-time, spoken conversation and the software emulation 238 of the simulator module 230 is programmed to mimic the software actually running a customer service center for which the trainee is training. For example, the screenshots, caller data format, options, keyboard configuration and the like are identical to that of the actual call center, so the trainee is trained on system operation at the same time the trainee is learning to communicate with customers (Koehler [0044]).

Further, Koehler demonstrates examples of both simulated and actual outputs in a voice application, wherein the use of a "nominal" output is addressed in an example, wherein Koehler teaches regional database may also include customer data, which likewise populates the software emulation 238, so that the trainee receives "pop-up" caller data during simulated calls, as would be the case in an actual call center interaction. The customer data includes, for example, the customer name, address, account balance, current product/service information and the like. The customer data may also be updated by a web interface (Koehler [0049]).

Furthermore, Koehler teaches that in response to the scenario initiation, the

business rules application 232 receives input from the trainee at step s324. The trainee's input includes speaking into the headset microphone or, alternatively, typing on a keyboard interfaced with the trainee terminal 130. Note that the initial dialog segment of the scenario includes an incoming customer call and the trainee's response, e.g., answering the call and greeting the customer. For example, the software emulation 238 may provide an indication of a queued call that mimics the same indication received by an agent at the actual call center, such as a beeping call initiation sound, indicating that the scenario is ready. The trainee receives the call (e.g., by a particular keystroke, as provided by the software emulation 238) and responds with a greeting (Koehler [0061]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Williams to incorporate deriving from the voice application a nominal output of the voice application as taught by Koehler to allow for voice files that may be referenced, prerecorded and stored as voice or text, having both text lines displayed and converted to speech (Koehler [0049]).

Re claims 21, 27, and 33, Williams teaches the method of claim 1, wherein the user simulation script is specified in a customized mark-up language (well known teachings [0010], user interface allows a user to graphically create a voice application, Fig. 4A, Fig. 8 test script generation).

Re claims 22, 28, and 34, Williams teaches the method of claim 1, wherein the step of processing further comprises simulating a text equivalent ([0050]) and an

execution time for each of the nominal output and the pre-determined user input.

However, Williams fails to teach an execution time for each of the nominal output and the pre-determined user input

Koehler teaches pre-determined time intervals contained text/speech segments that are to be outputted, wherein predetermined time intervals, a voice simulator operating with the simulator module 230 checks the processing queue table for new entries. When a new entry is identified, the voice simulator sends the phrase to a text-to-speech application to transform the text data into audio files (Koehler [0058]).

Koehler also teaches the identification of elapsed time used by an agent (i.e. agent in training), wherein the elapsed time 518 used to execute the dialog segment and the total coaching time 519 provided to the trainee (Koehler [0075] & Fig. 5, 518, dialog time).

Additionally, Koehler teaches actual and simulated data outputs, wherein Koehler teaches that simulator module 230 creates a realistic training environment in two respects. First, through the business rules application 232, the simulator module 230 initiates a real-time, spoken conversation and the software emulation 238 of the simulator module 230 is programmed to mimic the software actually running a customer service center for which the trainee is training. For example, the screenshots, caller data format, options, keyboard configuration and the like are identical to that of the actual call center, so the trainee is trained on system operation at the same time the trainee is learning to communicate with customers (Koehler [0044]).

Further, Koehler demonstrates examples of both simulated and actual outputs in

a voice application, wherein the use of a “nominal” output is addressed in an example, wherein Koehler teaches regional database may also include customer data, which likewise populates the software emulation 238, so that the trainee receives “pop-up” caller data during simulated calls, as would be the case in an actual call center interaction. The customer data includes, for example, the customer name, address, account balance, current product/service information and the like. The customer data may also be updated by a web interface. (Koehler [0049]).

Furthermore, Koehler teaches that in response to the scenario initiation, the business rules application 232 receives input from the trainee at step s324. The trainee's input includes speaking into the headset microphone or, alternatively, typing on a keyboard interfaced with the trainee terminal 130. Note that the initial dialog segment of the scenario includes an incoming customer call and the trainee's response, e.g., answering the call and greeting the customer. For example, the software emulation 238 may provide an indication of a queued call that mimics the same indication received by an agent at the actual call center, such as a beeping call initiation sound, indicating that the scenario is ready. The trainee receives the call (e.g., by a particular keystroke, as provided by the software emulation 238) and responds with a greeting (Koehler [0061]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Williams to incorporate an execution time for each of the nominal output and the pre-determined user input as taught by Koehler to allow for the mimicking or simulation of an interaction taking place within a call center (Koehler [0061]), wherein pre-determined time segments of speech/text can be output

(Koehler [0058]) to allow for the logging of a duration of a voice interaction (Koehler [0075] & Fig. 5, 518, dialog time)

Re claims 23, 29, and 35, Williams teaches the method of claim 1, wherein the simulated output simulates an output from a text to speech engine in response to the simulated input ([0050], & [0016] well known uses of TTS and speech recognition).

Re claims 24, 30, and 36, Williams teaches the method of claim 1, wherein the simulated output simulates an output from an automatic speech recognition engine in response to the simulated input ([0050], & [0016] well known uses of TTS and speech recognition).

Re claims 25, 31, and 37, Williams teaches the method of claim 1, wherein the simulated output simulates a pre-recorded audio source ([0049]).

Re claims 26 and 32, Williams teaches the method of claim 1, further comprising the steps of:

- a) deriving additional nominal outputs of the voice application;
- b) processing the user simulation script to ([0034]) generate additional simulated outputs ([0071]) for the voice application corresponding to the additional nominal outputs ([0047] & Fig. 6);
- c) processing the user simulation script to generate additional simulated inputs to the voice application ([0045]-[0047]); and

d) repeating steps a), b) and c) until the user simulation script is exhausted to simulate a complete set of user interactions with the voice application ([0050] & Fig. 6), in response to and as input for a complete set of user prompts from the voice application ([0066]-[0068]).

However, Williams fails to teach a) deriving additional nominal outputs of the voice application;

Koehler teaches a user application, for example a simulator module 230 causes options to be provided to the training terminal 130 at step s312. The options may be provided textually or audibly (Koehler [0052]). Further, in Fig. 4, Koehler teaches pre-determined responses in a voice application such, wherein voice files may be referenced, prerecorded and stored or produced on-line by the developer, for each of the desired customer lines in a dialog segment. The data for the new scenario is then stored in the database 212 as text. However, in an embodiment of the invention, the customer lines are converted to speech (unless they are already recorded speech) and also stored as audio files in .WAV or MP3 format, for example (Koehler [0087] & Fig. 4 element 414 for example).

Koehler teaches actual and simulated data outputs, wherein Koehler teaches that simulator module 230 creates a realistic training environment in two respects. First, through the business rules application 232, the simulator module 230 initiates a real-time, spoken conversation and the software emulation 238 of the simulator module 230 is programmed to mimic the software actually running a customer service center for

which the trainee is training. For example, the screenshots, caller data format, options, keyboard configuration and the like are identical to that of the actual call center, so the trainee is trained on system operation at the same time the trainee is learning to communicate with customers (Koehler [0044]).

Further, Koehler demonstrates examples of both simulated and actual outputs in a voice application, wherein the use of a "nominal" output is addressed in an example, wherein Koehler teaches regional database may also include customer data, which likewise populates the software emulation 238, so that the trainee receives "pop-up" caller data during simulated calls, as would be the case in an actual call center interaction. The customer data includes, for example, the customer name, address, account balance, current product/service information and the like. The customer data may also be updated by a web interface. (Koehler [0049]).

Furthermore, Koehler teaches that in response to the scenario initiation, the business rules application 232 receives input from the trainee at step s324. The trainee's input includes speaking into the headset microphone or, alternatively, typing on a keyboard interfaced with the trainee terminal 130. Note that the initial dialog segment of the scenario includes an incoming customer call and the trainee's response, e.g., answering the call and greeting the customer. For example, the software emulation 238 may provide an indication of a queued call that mimics the same indication received by an agent at the actual call center, such as a beeping call initiation sound, indicating that the scenario is ready. The trainee receives the call (e.g., by a particular keystroke, as provided by the software emulation 238) and responds with a greeting (Koehler [0061]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Williams to incorporate deriving from the voice application additional nominal output of the voice application as taught by Koehler to allow for voice files that may be referenced, prerecorded and stored as voice or text, having both text lines displayed and converted to speech (Koehler [0049]).

#### **(10) Response to Argument**

This is in response to the appeal brief filed 12/14/2009 appealing from the Office action mailed 07/15/2009.

#### **Re arguments directed to Williams in view of Koehler:**

- “Thus, despite the Examiner's assertions to the contrary (see discussion on page 6 of the Third Office Action), Appellants' use of the term is not "ambiguous." In contrast to a "nominal output," which is the voice application would otherwise provide to a user, a "simulated output" is an output in a simulation environment. Since, the actual output (i.e., nominal output) of a voice application would be a voice, the simulated output is text. Similarly, since the actual input into a voice application is audio, the simulated input is also text” (Appeal Brief page 10 last paragraph).**

After further consideration of the prior art of record, it appears that Williams in view of Koehler teaches “a simulated output for the voice application corresponding to the nominal output and a simulated input for the voice application corresponding to a pre-determined user input to the voice application”. Examiner has refined the grounds of rejection to direct the majority of the claim language to be taught by Williams as originally intended by Examiner, wherein Koehler now has been clarified to teach “deriving from the voice application a nominal output of the voice application”, whereby Examiner believes Koehler to teach user interaction for the voice application taught by Williams such as a simulation tool that gives an option to interact with text and voice.

Examiner acknowledges Appellants clarification during this appeal brief of what is intended to be claimed, such as “the actual output (i.e., nominal output) of a voice application would be a voice, the simulated output is text. Similarly, since the actual input into a voice application is audio, the simulated input is also text” (Appeal Brief page 10 last paragraph). However, that which is claimed does not reflect that defined in the arguments by Appellant (i.e. nominal output/input is voice and simulated output/input is text). This is in fact understood as speech to text.

In view of the above arguments by Appellant, Examiner now believes that Williams independently appears to teach both nominal output and simulated input, that is voice output and text input. Williams is directed to the use of a virtual call center capable of VXML using speech and text. For instance, Williams explicitly teaches that inputs to the tags are provided by analyzing the voice-capable markup language application to determine ***an expected voice input*** and ***a voice-recognized text input***

associated with a branch of the voice capable markup language application, and analyzing the expected **voice input** and the **voice-recognized text input** to determine an input to the tag (Williams [0070]).

Additionally Williams teaches that outputs of the tags are provided by **synthesized text-to-speech** ([Williams [0071]]). Thus Examiner strongly believes that Williams teaches “the actual output (i.e., nominal output) of a voice application would be a voice, the simulated output is text. Similarly, since the actual input into a voice application is audio, the simulated input is also text” as argued by Appellant, but more specifically, the claims language “a simulated output for the voice application corresponding to the nominal output and a simulated input for the voice application corresponding to a pre-determined user input to the voice application”.

Further, Williams teaches voice applications (pertinent to the present invention drawings Fig. 3, i.e. VXML browser and script), wherein one or more test scripts are generated in VXML software language. The one or more test **scripts are coupled to the test VXML server** 108, and are thereby executed within the test VXML server 108. When a step of the one or more test scripts is executed by the test VXML server 108, the execution causes a VXML test page to be called from the test VXML server 108. The VXML test page is coupled to the **VXML browser 110. The VXML browser 110 generates a simulated audio telephone signal** in response to the VXML test page and couples the simulated audio telephone signal to the telephony interface 112. As described above, the audio telephone signal can be either the signaling portion or the RT portion of a telephone call. The audio telephone signal is coupled to the PSTN,

whereby it is further coupled to the contact center 64. The contact center processes the audio telephone signal in the same way that the audio telephone signal of FIG. 2 is processed, thereby generating the IVR audio response onto the PSTN 62. The **IVR audio response is received** by the telephony interface 112, **whereupon it is converted to a response text and coupled to the VXML browser** 110 (Williams [0052-0053]).

Therefore, Williams explicitly teaches a simulated output for the voice application corresponding to the nominal output and a simulated input for the voice application corresponding to a pre-determined user input to the voice application.

With respect to Koehler, Examiner believes that Koehler provides a more advanced call flow representation of a voice application which not only handles text but displays it as shown in Fig. 4.

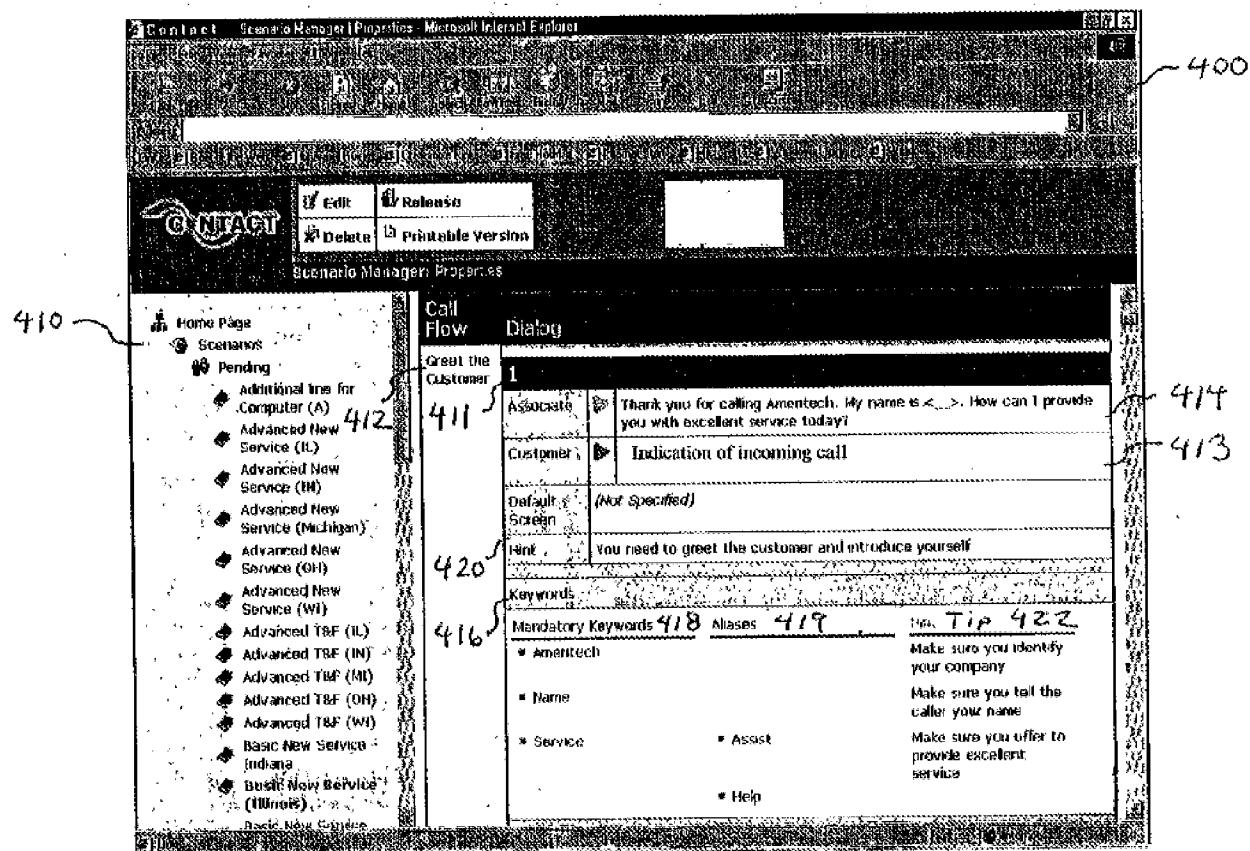


Fig. 4

Specifically Koehler improves Williams by teaching that when a new entry is identified, ***the voice simulator sends the phrase to a text-to-speech application to transform the text data into audio files***, which are stored, for example, in .WAV format, developed by Microsoft and IBM, or MP3 format. The audio files may be stored at the web sever 115 and thus retrievable by a universal resource locator (URL) or universal naming convention (UNC) path (Koehler [0058]).

Koehler teaches a user application, for example a simulator module 230 causes options to be provided to the training terminal 130 at step s312. The ***options may be provided textually or audibly*** (Koehler [0052]). Further, in Fig. 4, Koehler teaches

pre-determined responses in a voice application such, wherein voice files may be referenced, prerecorded and stored or produced on-line by the developer, for each of the desired customer lines in a dialog segment. The data for the new scenario is then ***stored in the database 212 as text***. However, in an embodiment of the invention, ***the customer lines are converted to speech (unless they are already recorded speech) and also stored as audio files in .WAV or MP3 format, for example*** (Koehler [0087]).

Therefore, Koehler improves the system of Williams to teach an interface that displays both text and speech during a voice application, allowing a user of a voice application to hear and read audio/text information.

#### **(11) Related Proceeding(s) Appendix**

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/Michael C Colucci/  
Examiner, Art Unit 2626

/Richemond Dorvil/  
Supervisory Patent Examiner, Art Unit 2626

**Conferees:**

Richemond Dorvil  
/R. D./  
Supervisory Patent Examiner, Art Unit 2626

/Talivaldis Ivars Smits/  
Primary Examiner, Art Unit 2626